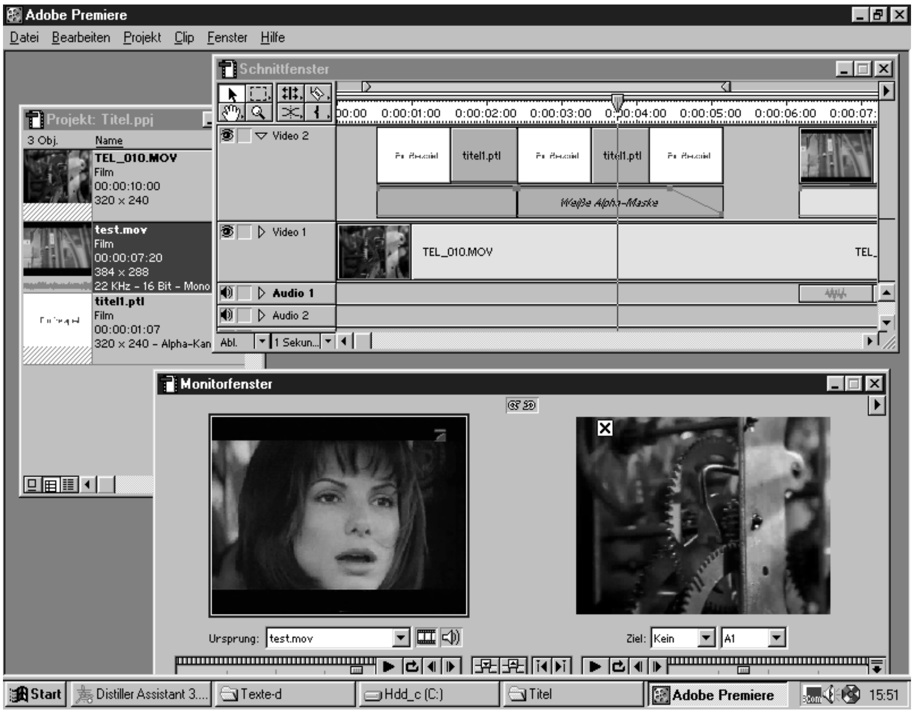
Video capture cards have become so cheap that a PC can now be used for sophisticated video editing. Nearly all the cutting techniques and effects that we commonly see on TV can be achieved with a video capture card in combination with suitable software: blends and fades, digital effects, filters, animations, titling and so on. Of course, this all requires a lot of disk space, and calculating the various effects is very time-consuming. Practically speaking, this means that video clips prepared using an inexpensive editing system with 'normal' hard disks should be limited to less than 10 minutes of playing time.

By Dr M.Seiwert

Video processing on the PC

Part 1: basics



Video editing systems, consisting of a video capture card and matching video editing software, are available starting at around £200. The video capture card must be installed in a free PCI slot that has bus master capability, so that it can directly transfer data to and from the memory without the intervention of the processor. Video capture cards have VHS inputs and outputs with Cinch jacks, and most of them also have Super-VHS inputs and outputs with mini-DIN jacks. In use, the card is connected to a camcorder or video recorder, which acts initially as a playback source. It is preferable to use a S-VHS connection, since it transfers the chrominance and luminance signals separately. This reduces the crosstalk between the two signals. The sound signal is either processed in a separate part of the video capture card or connected directly to a sound card, depending on the particular video capture card used. The sound signal is connected via a Cinch cable. **Figure 1** shows the block diagram of a typical video capture card.

When the video capture card receives a television signal (from a video recorder, for example), it decodes the input signal, digitizes it and converts it into a sequence of bit-mapped images. Depending on the design of the video capture card, the images are displayed on the computer monitor and/or a television set that serves as a control monitor. Each bit-mapped image occupies 1.4 MB at full resolution, which is roughly the capacity of a 3.5-inch diskette. The computer must of course handle up to 30 such images per second and store them in memory. Since even fast disks cannot handle data at this rate, the bit-mapped images are compressed before they are stored. Special ICs on the video capture card are used to perform this compression at the real-time image data rate, since the PC processor cannot normally handle this task. The compression ratio can be adjusted for the best compromise between the disk data transfer rate and image quality; this depends on the type and configuration of the PC used. If the compression ratio is too small, the disk will not be able to keep up and image data will be lost. On the other hand, if the compression ratio is unnecessarily high then the image quality will be degraded by compression artifacts.

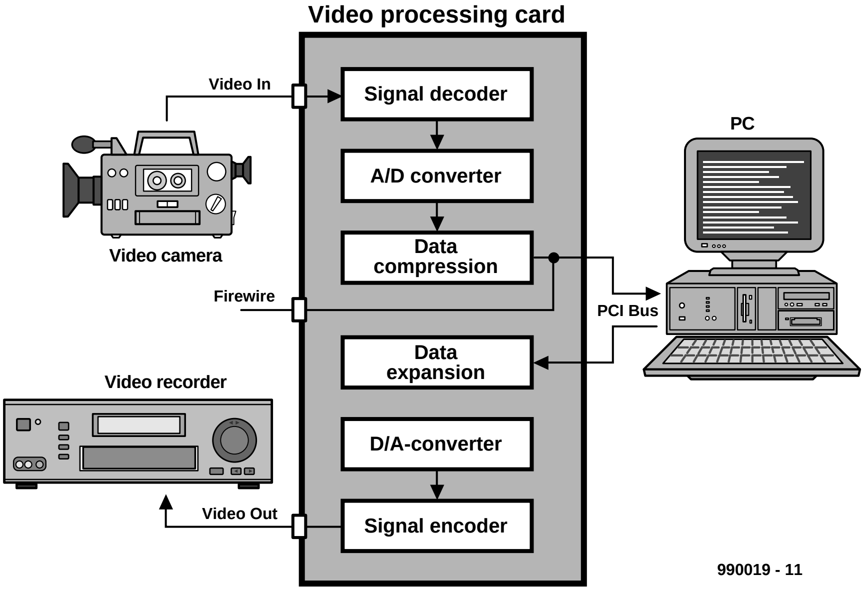


Figure 1. Block diagram of a typical video capture card.

The compressed image data are stored together with the sound data as *.avi files on the hard disk. These AVI files are different from the AVI files that you may be familiar with from multimedia applications in one important aspect: they can only be used on the system where they were generated, or on a system with the same model of video capture card. For playback, the image data must be decompressed at the required rate, which also takes place on the video capture card.

Many video capture cards display played-back AVI files only on a television, and not on the computer monitor. The AVI files are the starting point for video editing, which is discussed in part 2 of this article. Once a video clip has been fully edited, it can be

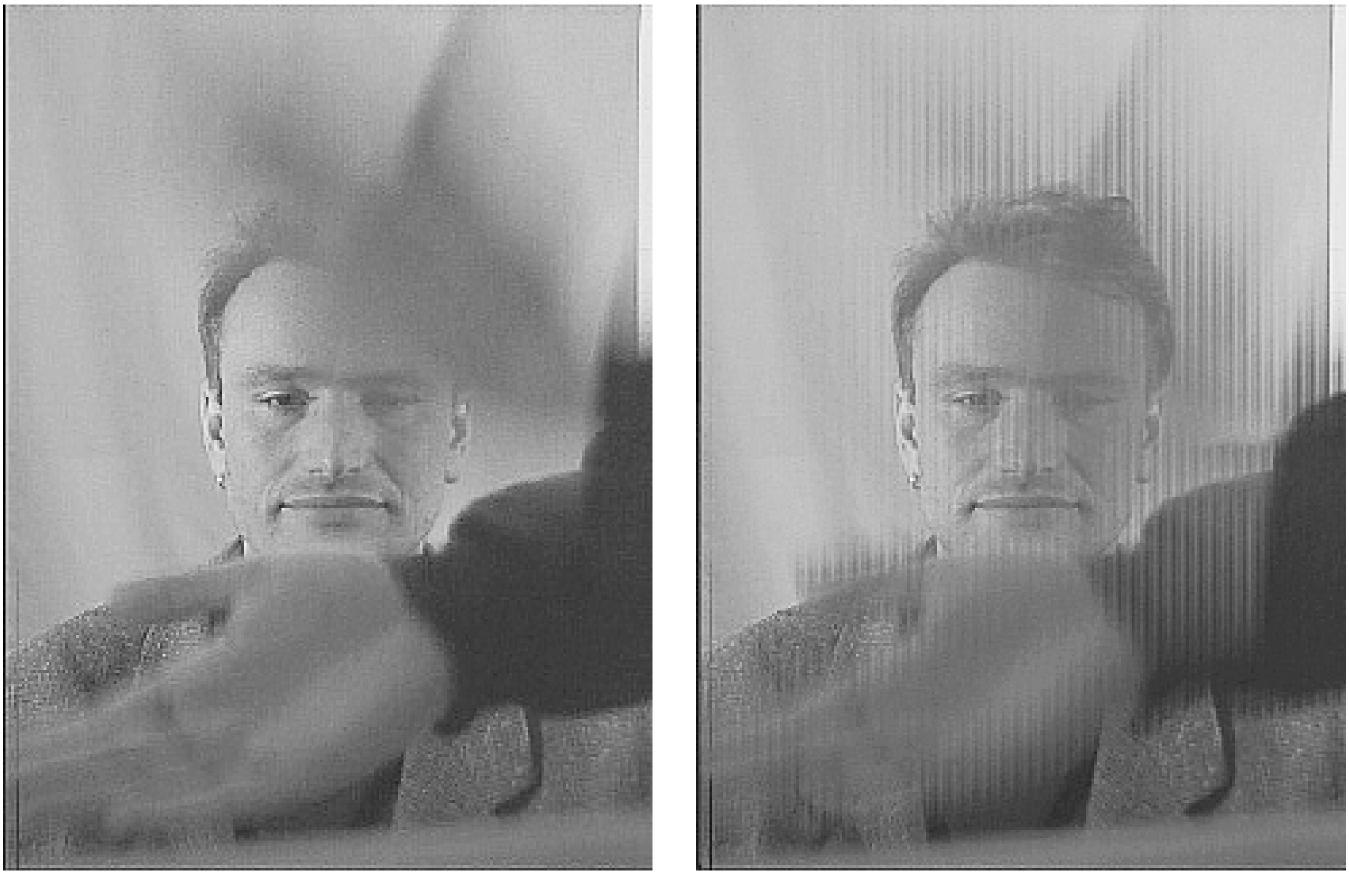


Figure 2. An example of the comb effect. In this recording of a speech, the missing clappercard was replaced by a handclap. Left: no comb effect, right: comb effect

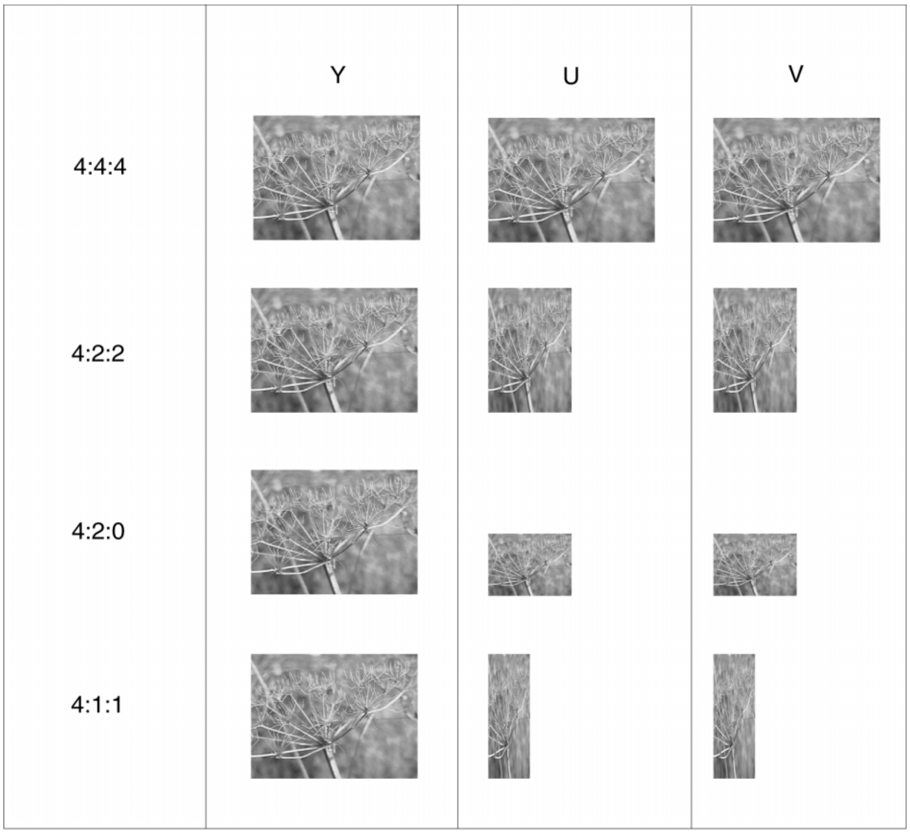


Figure 3. Summary of commonly used black/white and colour resolution relationships.

‘played back’ to the video recorder. The outputs of the video capture card are connected to the inputs of the video recorder for this. Playing back an AVI file involves the same processes as capturing a video signal, in the reverse order. The digital video and sound data are decompressed by the video capture card and converted into an analogue television signal that appears at the video output connectors. This signal can be recorded by the video recorder.

Alternatively, the finished clip can be further processed to produce AVI files that can be played back on any standard PC. This requires the data compression and decompression to be performed by a software algorithm, instead of by the codec hardware of the video capture card. The algorithm is independent of the video capture card and is executed by the PC’s processor alone. Naturally enough, AVI files prepared in this manner are displayed on the computer monitor.

Table 1. The relationship between the compression ratio, data rate, file size and image quality.

Compression-ratio	Data rate, MB/s	Duration per GB, min:sec (PAL)	Image quality and applications
3:1	7	2:26	Digital Betacam – television, advertising
5:1	3,5	5:16	DV – news reporting, digital computer videos
7:1	3	5:41	S-VHS, Hi-8 – news reporting, analog computer videos
10:1	2,1	8:08	S-VHS, Hi-8 – computervideos
15:1	1,4	12:11	VHS – amateur projects

Digital camcorders and video capture cards

Digital camcorders have been commercially available for some time now. They are wonderfully compact, and their picture quality (relative to the demands of the consumer market) is excellent. In a digital camcorder, the video signal is digitized and compressed before it leaves the camcorder. The tape drive records the compressed digital signal. This means that the first three of the video signal processing functions that usually take place in a video capture card have already been performed in the camcorder.

There are ‘digital’ video capture cards that are designed to work with digital camcorders. These have connectors for the digital Firewire bus, in addition to the usual VHS and S-VHS connectors. (For an explanation of Firewire, see the PC Topics Supplement in the February 1998 issue of *Elektor Electronics*.) This allows the already digitized camcorder data to be transferred directly from the video cassette to computer files. The AVI file format is not used in this case; instead, the ‘digital video’ format is used to produce files with a *.dv extension. This file format directly corresponds to the data format on the video cassette. The compression ratio is not selectable in this case. Instead, it is fixed, and thus a sustained data transfer rate of 3.5 MB/s is required. If the hard disk drive cannot support this data rate, it’s time for an upgrade. Of course, these cards can also capture and store standard video signals via their analogue inputs. Such signals are digitized, compressed and stored in the DV format, in exactly the same manner as a digital camcorder works when making a video recording.

Special aspects of the video signal

There are several different video signal standards. NTSC is commonly used in the US and Japan, while PAL and SECAM are used in Europe. This is not a problem, since the video capture cards can handle all these standards, and they can to a certain extent even automatically detect which standard is being used. Nevertheless, we as users must be aware of the major differences between the various standards, so that we can make the best selections for the capture parameters. NTSC uses a field rate of 60 Hz, which means that 60 fields (rasters) or 30 full frames (‘pictures’) are transmitted per second. In Europe, where the standard

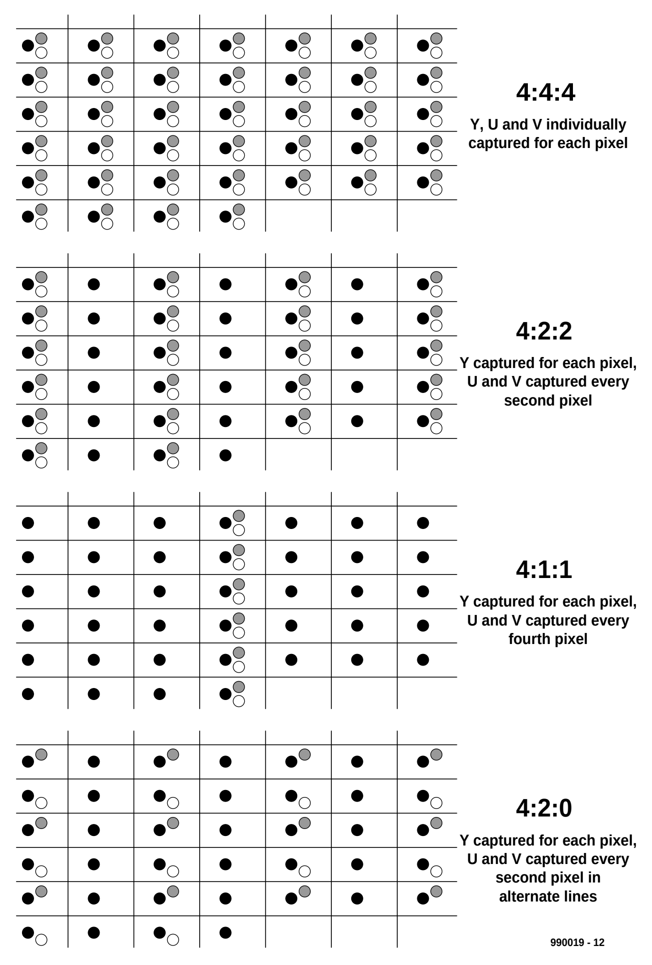
Sampling rates for digital video formats

The compleet video signal, consisting of a luminance component (Y) and two chrominance components, U and V, is stored as three monochrome (black and white) components. When all picture elements (pixels) of these three monochrome images are captured (sampled), the procedure is called 4:4:4 recording (top diagram). Because of the immense amount of data generated in this way, the system is rarely used in practice.

When 4:2:3 recording is applied, the U and V colour components are only sampled every other pixel. This results is a 1/3 reduction in the amount of data required for 4:4:4 recording.

With 4:1:1 recording, every fourth pixel is sampled in the colour channels. The data compression factor with respect to 4:4:4 recording is 50%.

As an alternative to 4:1:1 recording, a system called 4:2:0 is generally used in Europe. Only one colour component is sampled on every second pixel, the sampling being alternate on the U and Y components.


4:4:4
Y, U and V individually captured for each pixel
4:2:2
Y captured for each pixel, U and V captured every second pixel
4:1:1
Y captured for each pixel, U and V captured every fourth pixel
4:2:0
Y captured for each pixel, U and V captured every second pixel in alternate lines
990019 - 12

A comparison between different sampling rates. Solid dots represent brightness information (luminance, Y), the open dots and grey dots represent the colour components U and V respectively.

mains voltage frequency is 50 Hz, the rate is 50 fields per second or 25 frames per second. It is thus a good idea to capture NTSC video at 30 frames per second (fps) and PAL or SECAM video at 25 fps.

The picture dimensions are also different: NTSC video converts to bitmap images of 640 ⇔ 480 pixels at full resolution, while PAL and SECAM produce 768 ⇔ 576 pixels at full resolution. The vertical and/or horizontal resolution for data capture can be halved in order to reduce the size of the image data files. For example, if you want to produce computer videos with a resolution of 320 ⇔ 240 pixels, you can capture the video data at half of the full vertical and horizontal resolution without affecting the final image quality. The resulting data files will use only one fourth as much disk space as full-resolution files.

Half frames and the comb effect

In order to keep flickering to an acceptable level, television pictures at 25 or 30 frames per second are transmitted as alternating even and odd half frames at 50 or 60 fields per second. The first field (the even half frame) contains all even-numbered lines of the full frame, and the second field (the odd half frame) contains all odd-numbered lines (although this sequence may be reversed, since it is not standardized). This means that the picture that you see is refreshed 50 or 60 times per second, which significantly reduces the flickering effect. Video capture cards play along with this trick, and in fact capture one half frame every 50th of a second if they are configured for a 25 fps acquisition rate. In exactly the same way, the card generates 50 half frames per second, which are faithfully registered by the video recorder. If you use the system only for editing video clips, everything is in order and you need not be especially concerned about frames and half frames. The only thing that you might notice is that stationary ('halted') pictures from action scenes will flicker quite visibly on the control monitor. This is due to the fact that the video capture card displays both half frames in rapid sequence on the control monitor, even when playback is halted. If the two half frames are not identical, due to rapid movements, there will be visible flickering.

Problems can crop up if you want to capture video images for computer playback. PCs do not have enough processing power to update the monitor display 50 or 60 times per second. They therefore combine successive half frames into single composite frames. This violates the time continuity of the video imagery, since half frames that were recorded at two different times are displayed at the same time. This causes the image to be disrupted if there is rapid motion, This is called the comb effect because of the way it affects the picture (see **Figure 2**).

There is a simple way to avoid the comb effect: capture the video at one half of the full vertical resolution. In this case only every second raster line is captured, which is the same as capturing only every second half frame. The resulting 'reduced' resolution of 288 lines is adequate for computer videos, given the current level of PC technology. This number of lines also corresponds to the MPEG-1 standard.

Data reduction

We have no choice but to reduce the volume of the video data, so that the hard disk can handle the data stream and the file sizes remain manageable.

As a first measure, we exploit a characteristic of human visual perception that has long been known from photography and television. We perceive detail at much higher resolution for the monochrome (black-and-white) component of an image than for the colour components. Since a composite colour television signal consists of a brightness component (containing the monochrome imagery) and two colour components (U and V), an obvious approach is to store the colour components at reduced resolution. Our eyes, which are 'half blind' in the colour realm, hardly notice the reduced resolution. **Figure 3** lists commonly-used resolution relationships for video data storage and shows sample images for each one.

For consumer applications, the relationships 4:2:0 (for PAL and SECAM) and 4:1:1 (for NTSC) are generally used. Here the colour components are stored at only one fourth the resolution of the black and white component. This is also considered acceptable in the professional realm, such as television news reporting. For costly advertising productions and other productions that require a certain quality margin, due to extensive digital manipulation of the imagery, a relationship if 4:2:2 is employed.

In addition to this reduction in the resolution of the colour components, the individual images are compressed. This is done by applying a discrete cosine transformation (DCT) followed by quantization. The popular JPEG image format, for example, is based on DCT processing. This technique is based on the fact that in a 'real' image the brightness and colour values vary only slightly in the neighbourhood of any given pixel, except at the edges of the image. This 'low-pass' character of natural images means that we can safely ignore the high-frequency components.

The compression ratio is adjustable with analogue video capture cards, so that it can be matched to the capability of the hard disk drive. The processing speed of the video capture card is also important, since the minimum level of compression that the card can handle determines the maximum quality that is possible for the captured video imagery. Table 1 clarifies the relationship between image quality and the compression ratio.

(990019-1)